

Age-Related Differences in the Perception of Eye-Gaze from a Social Robot

Lucas Morillo-Mendez¹[0000-0001-7339-8118], Martien G.S. Schrooten²[0000-0002-9462-0256], Amy Loutfi¹[0000-0002-3122-693X], and Oscar Martinez Mozos¹[0000-0002-3908-4921]

Important Note: This paper is a preprint of an upcoming publication to appear in the Proceedings of the 13th International Conference on Social Robotics (ICSR 2021): <https://www.colips.org/conferences/icsr2021/wp/>

Abstract. The sensibility to deictic gaze declines naturally with age and often results in reduced social perception. Thus, the increasing efforts in developing social robots that assist older adults during daily life tasks need to consider the effects of aging. In this context, as non-verbal cues such as deictic gaze are important in natural communication in human-robot interaction, this paper investigates the performance of older adults, as compared to younger adults, during a controlled, online (visual search) task inspired by daily life activities, while assisted by a social robot. This paper also examines age-related differences in social perception. Our results showed a significant facilitation effect of head movement representing deictic gaze from a Pepper robot on task performance. This facilitation effect was not significantly different between the age groups. However, social perception of the robot was less influenced by its deictic gaze behavior in older adults, as compared to younger adults. This line of research may ultimately help informing the design of adaptive non-verbal cues from social robots for a wide range of end users.

Keywords: Human-Robot Interaction, Older Adults, Non-Verbal Cues

1 Introduction

In the last years, there has been an increasing interest in the use of social robots to assist older adults (OA) during daily life tasks [20]. An important cue in the interaction with social robots is non-verbal communication such as deictic gaze [17,2]. Humans use deictic gaze to guide the attention of another person towards a point in the space by looking at it. This communicative signal is key to initiate a shared attention between individuals and to increase the efficiency in collaborative tasks [4]. In addition, deictic gaze is important in OA because it can help to inform age-related differences in human-robot interaction (HRI) [7]. This is because the sensibility to deictic gaze declines naturally with age, reflecting a reduction in social perception in OA [21]. For this reason, it is important to explore how deictic gaze is attended in normal aging when performed by a robot.

At the same time, there is a call for more studies regarding non-verbal cues in which OA are direct research participants, in contrast to studies where OA

act only as beneficiaries, and to further compare the outputs with younger controls [22]. Therefore, it is relevant to explore the benefits of non-verbal cues from social robots towards OA during collaborative daily life tasks, and how age-related differences may influence their perception of a social robot in contrast to younger populations. These studies may help improving the design of non-verbal cues in HRI that adapt to age changes.

This work seeks to explore potential age-related differences in the perception of deictic gaze from a social robot when collaborating in tasks inspired by daily life activities. To do so, we designed a set of online visual search tasks with a video recording of a Pepper robot¹ given its wide use in research related to HRI. Pepper does not have degrees of freedom in the eyes to reflect human-like gaze. Therefore, and in line with previous research [2,18], we used its head movement to point to objects as a way to reflect deictic gaze as shown in Fig. 1.

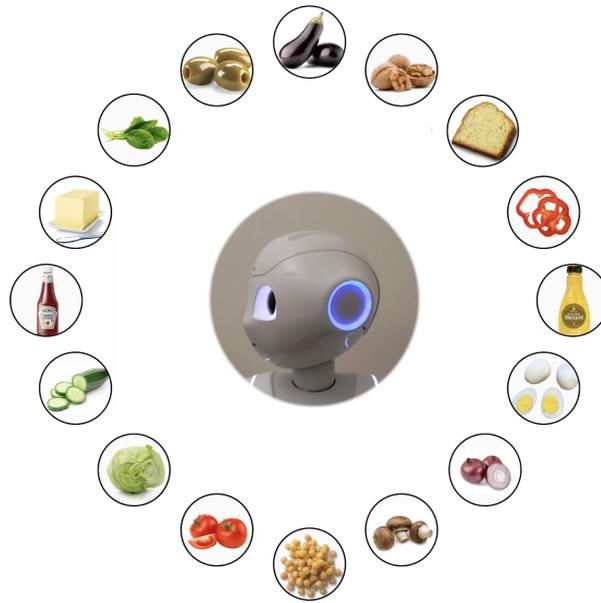


Fig. 1. Layout of the task. The video of the robot provides verbal instructions switching between static positioning and deictic gaze behaviour. The picture shows a trial in which Pepper uses deictic gaze towards the ketchup, where participants should click.

¹ <https://www.softbankrobotics.com>

2 Related Work

Gaze behavior from a robot is highly studied in HRI. It has been shown that making eye contact with a robot evokes similar physiological responses as if with a human [13]. Appropriate gaze from a speaking robot towards a human can positively affect the recall of what has been said [16] and can regulate the role of the participants in conversations [17]. Similarly, a robot moving the head away can effectively reflect gaze aversion and can be perceived as more thoughtful by the users [3]. In collaborative scenarios where a human follows guidance from a robot, deictic gaze from the robot can assist the human partner by signaling at objects in space [1], although the specific benefits differ among studies [2,18,14].

The work in [14] investigated deictic gaze in a situated human-agent collaboration. The results showed that this non-verbal cue led to higher interaction times and, thus, inferior task performance. In contrast, the work in [18] found that deictic gaze in the form of head movement from a robot did not affect task-completion times, although it helped to reduce the number of errors. This is in line with [2], which suggested that deictic gaze from a robot is not that useful in simple tasks when compared to difficult ones.

Previous research also indicated that eye-gaze following deteriorates with age [21]. Nevertheless, to the best of our knowledge, the influence that gaze from a robot may have on OA has not been explored. In our study we used a similar task as in [14] reflecting a realistic interaction between OA and an assistive robot. Although the remote nature of our study might limit the interaction between the users and the robot, the current approach, importantly, allowed us to control the influence of some extraneous variables on the main outcome variables, such as the social presence caused by the robot looking at the user, which may lead the users to start a conversation with the robot and get higher completion times as reported in [14].

3 Methods

This study was performed during the Covid-19 global pandemic. We designed an experimentation method for effective remote participation, which also facilitates larger scale testing. More specifically, we designed a controlled online collaboration task mimicking an everyday life situation (see [15] for a description of an equivalent face-to-face interaction). In this experiment, a video of Pepper verbally guided the participants during a task that represented a guided preparation of a sandwich recipe. To compare the potential benefit in the perception of non-verbal cues, the robot switched its behaviour between static-based and deictic gaze-based indications. We measured the participant’s reaction times and task-completion times during task performance. An example video of the task is available².

² <https://youtu.be/6zSgm8jEnCM>

3.1 Scenario

Our experimental scenario featured a video of a Pepper robot who verbally guided participants through two everyday-like visual search tasks which consisted on clicking on several ingredients for preparing a sandwich. An example layout of the ingredients and the robot is shown in Fig.1. We defined two robot conditions: a *static robot* (SR) which always looked at the camera while giving instructions, and a *moving robot* (MR) which also featured deictic gaze by moving the head towards the correct ingredient. Each of these conditions defined a task: a SR task, and a MR task. Inside each task, the user had to prepare two sandwiches by following the verbal instructions of the Pepper robot, which named each ingredient and waited for the user to click on it. So, each participant prepared two consecutive sandwiches with a static robot (SR), and two other consecutive sandwiches with a moving robot (MR). The full structure of the experiment is shown in Fig. 2.

A trial consisted of the selection of one ingredient. The user prepared two sandwiches from five ingredients in each task. The order of the sandwiches within a task and the order of the ingredients in each sandwich were fixed. The order of the tasks was counter-balanced and started randomly either with SR or MR.

Within each task, we measured reaction time (RT) and task-completion time (TCT). We defined RT as the time span between the moment the robot started naming one ingredient and the moment the participant clicked on that ingredient within a trial. The RT of the *bread* was excluded because of its high predictability (always first/last ingredient). A trial was correct when the mentioned ingredient was clicked. The final number of trials per task in which RT was used was ten: five ingredients for each sandwich. We defined TCT as the time needed by a participant to finish one task (two sandwiches).

3.2 Materials

After each task, participants were presented with a set of self-report questionnaires and subscales. First, the mental demand subscale of the NASA-Task Load Index (NASA-TLX) [12] was used for assessing the mental demand between robot conditions. Second, the Godspeed Questionnaire Series [5] was used to measure the perceived anthropomorphism of the robot by the user. Here, we used a modified version of the anthropomorphism questionnaire due to an irrelevant item for the context of this study (*moving rigidly-elegantly*). Moreover, we added the item *mechanical-organic* as in [6]. Finally, the RoSAS [6] scale was used to measure the perception of warmth, competence, and discomfort caused by a robot. We added two extra questions at the end: Q1) *Did you notice any difference between the robots in the tasks?* to check whether the person was aware of the difference between robot conditions; and Q2) *Which robot did you prefer from the ones you interacted with?* to check their preferred condition (SR or MR).

This study was built using *Labvanced* [10], an online tool for designing and remotely distributing experiments on human cognition. The language used for the whole study was Spanish.

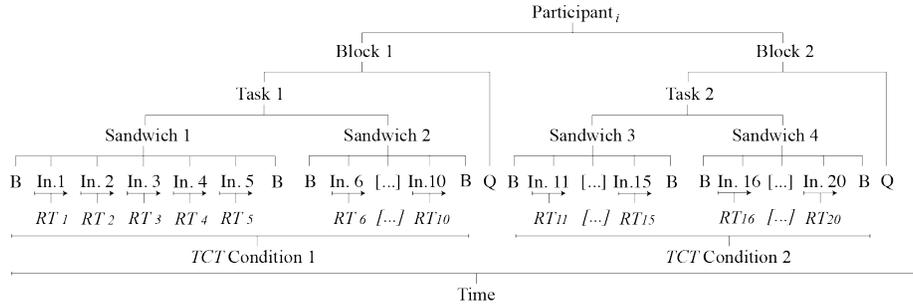


Fig. 2. Structure of the experiment for a participant i . The letter ‘B’ represents the bread, while ‘In. x ’ represents ingredient in position x . ‘Q’ refers to the questionnaire.

3.3 Experimental Design

Our study followed a 2x2 mixed design with two robot conditions, SR and MR (within-subject), and two age groups, Adults (A) and OA (between-subjects). The age range in the A was ≥ 18 and < 65 years, and in OA it was ≥ 65 . This division was based on the working retirement age.

The presentation order of the blocks (see Fig. 2) was counter-balanced and participants were randomly assigned to one of the two possible orders. We explored age-related differences in the effect of the robot’s deictic gaze on task performance and social perception of the robot, as reflected in the interaction between age group and robot condition

3.4 Sample

We performed a G*Power analysis [8] to calculate a minimum sample size that allowed an expected power $(1 - \beta)$ of 0.80 to detect a small effect size of $f = 0.25$ ($\eta^2 = 0.06$) between age groups. The result of the analysis indicated a minimum required sample size of 98 participants. A total of 329 participants took part in the study, of which 53 were excluded due to incomplete data. A summary of the characteristics of the final sample is shown in Table 1.

Potential participants were contacted via mailing lists from Spanish universities with adult education programs. Inclusion criteria (based on self-report) were to be fluent in Spanish, to have normal or corrected-to-normal vision, and to be cognitively healthy. Participants gave written informed consent in accordance with the Declaration of Helsinki and were informed about research goals. Participation was voluntary and no personal data that allowed their identification were obtained. In addition, approval was obtained from the corresponding program coordinators at each university.

The age range was 18 – 64 for A and 65 – 88 for OA. The mean age between the groups was significantly different ($t(145) = 14, p < 0.001$). A chi-squared test showed no significant differences between the age groups in their level of education, their previous knowledge of Pepper, and their experience with computers.

Table 1. Final sample description

Group	Age (years)		Gender			Comfort w/ computers			Had seen Pepper before		
	Mean	SD	M	F	Other	No	Not Sure	Yes	No	Not Sure	Yes
OA	69.3	3.8	76	74	0	1	7	142	108	26	16
A	53.4	12.1	45	80	1	1	11	113	84	24	18

Whereas there were about as many men as women in the OA group (51% men; 49% women), men were underrepresented in the A group (35.7% men; 63.5% women), ($\chi^2(1) = 5.9, p = 0.014$).

3.5 Procedure

The experiment started by asking participants to put on headphones (in order to reduce potential external noise) and to calibrate the volume of their headphones to ensure they could hear the robot clearly. Then, they were informed in the consent form about the study and the possibility of ending it at any time and filled a sample information questionnaire. Before the experiment started, participants had time to get familiar with the interface and ingredients to be used in the task. This last step was done to reduce the possibility of a poor performance derived from a participant not knowing an ingredient. To reduce external influences, participants were encouraged to avoid distractions and to be rested before starting. To favour this, we kept the experiment short and they were informed about its duration, fifteen minutes. For the main tasks, they were also encouraged to perform as well as they could, but without explicit instructions about being fast. This was to maintain the everyday nature of the task in contrast to a classic computerized experiment. In addition, they were not warned about the difference between the robot conditions (SR or MR).

4 Results

4.1 Reaction Times and Task-Completion Times

We first present reaction times (RT) and task-completion times (TCT) between the robot conditions (SR, MR), age groups (OA, A), and the combination of both. Potential noise in these time measures from participants due to external factors such as computer, browser, or operative system, was corrected using metadata provided by Labvanced. To analyze the RT data, we used the median RT of the correct trials within each task per participant (Fig. 2). The percentage of incorrect, and thus excluded, trials was 2.24%. Due to violations of assumptions for the mixed ANOVA test, we analyzed the data using a Mixed Robust ANOVA test with 20% trimmed means and 2000 bootstrapped samples.

The RT for different age groups and conditions are shown in Fig. 3A (left). Means, standard deviations and p-values for RT are also reported in Table 2.

Values were significantly different between the robot conditions, showing a facilitation effect for the MR condition. We also found a main effect of age that showed higher RT for OA. We did not find an interaction effect between robot condition and age group. Following [21], the strength of the facilitation effect was calculated as a proportional difference score $(RT_{SR} - RT_{MR})/RT_{MR}$. An independent robust t-test (trim = 0.2, samples = 2000) did not show significant differences in the strength of the facilitation effect caused by the MR in RT between age groups (Fig. 3B and Table 2), i.e., the magnitude of the help provided by the MR in terms of how fast the ingredients were clicked was similar between the age groups.

The TCT for different age groups and conditions are shown in Fig. 3A (right). Means, standard deviations and p-values for TCT are also reported in Table 2. Results show a facilitation effect for the MR condition in TCT. Furthermore, TCT was also higher for OA. We could not find any interaction effects between robot condition and age group. The strength of the facilitation effect was also calculated as a proportional difference score $(TCT_{SR} - TCT_{MR})/TCT_{MR}$. A robust t-test did not show significant differences in the strength of the facilitation effect caused by the MR in TCT between age groups (Fig. 3B and Table 2). Finally, the age groups did not differ in how fast they performed the task.

4.2 Questionnaire Measures

We now present the scores of the different questionnaires and scales from Sect. 3.2. We used a Robust Mixed ANOVA for all the scores except for anthropomorphism, as it met the assumptions for a regular Mixed ANOVA. Table 3 shows the social perception scores and the Cronbachs's α of each construct with the corresponding analyses. We found a significant effect of age by which OA perceived the robots as more anthropomorphic as compared to A. In addition, the MR scored significantly higher than SR in all the social perception scores except in the discomfort score. This indicates a more positive perception of the MR. Finally, a significant interaction effect showed that the deictic behavior of the robot had a lower impact on the self-report of anthropomorphism and discomfort in the OA group.

To test whether the deictic gaze of the robot affected the reaction times even if the participants were unaware of that movement, we analyzed the subset of all participants who retrospectively reported to not have noticed the difference between the robot conditions by answering *No* to Q1 (see Sect.3.2).

In the subset of participants who answered *No* to Q1 (a total of 116), we found (1) a main effect of age group on RT at $p < 0.001$, but no effect of robot condition or interaction effect; (2) a main effect of age group on TCT at $p < 0.001$, but no main effect of robot condition or interaction effect; (3) no effects in any of the subjective scores; (4) an over-representation of OA (67.2%), as compared to A (32.7%) ($\chi^2(1) = 12.5, p < 0.001$).

We also analyzed the participants who expressed a preference for a robot in their answer to Q2. From a total of 163 answers, 78.5% chose the MR. For the participants choosing the MR, the differences between age groups, A=57.8%,

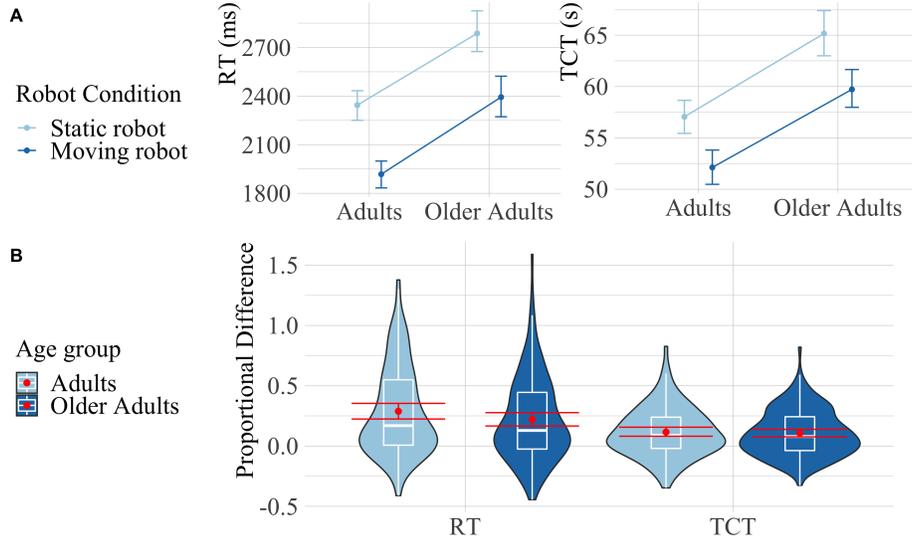


Fig. 3. A) Mean Reaction Time (left) and Task-Completion Time (right) for each group. B) Violin plots with means in red of the proportional differences between robots. Error bars show 95% bootstrapped confidence intervals.

OA=42.1%, were not significant ($\chi^2(1) = 2.47, p = 0.11$). Finally, with respect to the mental demand scale, we found no main effect of age ($p = 0.07$) or robot condition ($p = 0.11$), and no interaction ($p = 0.33$) ($M = 4.18 \pm 3.75$).

5 Discussion

This work sought to explore potential age-related differences in the perception of visual cues from a social robot. We focused on the influence of deictic gaze during a collaborative task inspired by daily life activities. We found a facilitation effect of deictic gaze from a Pepper robot in all the participants independently of their age. Given this facilitation effect for both time scales, our main interest was to find if its magnitude was different between age groups. Our results showed that the facilitation effect from the deictic gaze was not significantly different between age groups, neither in TCT nor RT (Fig. 3; Table 2). To further investigate the effect of deictic gaze, future research could include additional control conditions like a human face or non-social signalling. In addition, the high predictability, and thus the potentially high trust placed in a robot who always signals the correct ingredient (i.e., 100% valid gaze cues), could have had an influence on the speed of the responses. Future studies might want to also include invalid gaze cues that signal the incorrect ingredient and neutral gaze cues that are not informative about the location of the ingredients [21]. Comparing the impact of valid, invalid, and neutral gaze cues would help to explore the impact of credibility of the robot and trust in the system on task performance, as well as possible attentional costs [22] as they occur with human gaze cues.

Table 2. Means, SDs and main/interaction effects on the times

	RT (ms)	TCT (s)	% Facilitation (RT)	% Facilitation (TCT)
Age	$p < 0.001$ (***)	$p < 0.001$ (***)	$p = 0.31$	$p = 0.66$
M^A	2131 ± 575	53.5 ± 9.8	0.29 ± 0.36	0.12 ± 0.2
M^{OA}	2590 ± 808	61 ± 12.5	0.22 ± 0.35	0.11 ± 0.2
Robot	$p < 0.001$ (***)	$p < 0.001$ (***)	-	-
M^{SR}	2585 ± 721	61.5 ± 12.3		
M^{MR}	2176 ± 716	56.3 ± 11.1		
Age*Robot	$p = 0.2$	$p = 0.7$	-	-
M^{SR-MR}	426 ± 582	4.9 ± 10.7		
M_{OA}^{SR-MR}	393 ± 744	5.4 ± 11		

Table 3. Means, SDs, Cronbachs's α and main/interaction effects on the social perception scores

	Anth. ($\alpha = .88$).	Warmth ($\alpha = .87$)	Compt. ($\alpha = .84$)	Discom. ($\alpha = .76$)
Age	$p = 0.02$ (*)	$p = 0.051$	$p = 0.26$	$p = 0.88$
M^A	2.66 ± 0.94	2.47 ± 0.8	3.64 ± 0.67	1.71 ± 0.58
M^{OA}	2.89 ± 0.77	2.56 ± 0.8	3.65 ± 0.73	1.73 ± 0.58
Robot	$p < 0.001$ (***)	$p = 0.003$ (**)	$p < 0.001$ (***)	$p < 0.001$ (***)
M^{SR}	2.7 ± 0.92	2.42 ± 0.8	3.49 ± 0.76	1.76 ± 0.6
M^{MR}	2.87 ± 0.7	2.62 ± 0.79	3.8 ± 0.61	1.68 ± 0.55
Age*Robot	$p = 0.03$ (*)	$p = 0.85$	$p = 0.47$	$p = 0.003$ (**)
M^{SR-MR}	-0.27 ± 0.77	-0.25 ± 0.67	-0.39 ± 0.71	0.173 ± 0.58
M_{OA}^{SR-MR}	-0.08 ± 0.66	-0.15 ± 0.62	-0.25 ± 0.64	0 ± 0.5

OA scored higher in anthropomorphism regardless of the robot condition. There was also an increase in all the social perception scores as a result of the robot deictic gaze (Table 3). Moreover, the MR was chosen as favourite. These results support previous notions that appropriate social behaviors improve the acceptability of social robots [9]. All the participants indicated a low mental demand ($M = 4.18$ out of 21) when performing the tasks, independently of the robot condition. Moreover, we found an interaction effect in the scores of anthropomorphism and discomfort caused by the robot (Table 3) which varied less between robot conditions in OA. This interaction suggests a different perception of deictic gaze from a robot by OA.

A proportion of 42% of the participants reported not detecting the differences between the robot conditions. For this group we only found age effects on RT and TCT (Sec. 4). It cannot be excluded that participants interpreted Q1 (Sect. 3.2) literally and therefore reported no physical differences between the robots. However, this subgroup did not show a facilitation effect of the deictic gaze. Despite these participants not showing a facilitation effect, the incorrect trial ratio in our sample remained low (2.24%). Notably, within this subgroup, there were more OA than A. This is in line with previous work [21] showing the age-related decline in eye-gaze following. Alternatively, the over-representation of OA in this subgroup might as well reflect a broader cognitive decline, or difficulty in remembering the differences between robot conditions [11].

There are two main limitations in this study that could be addressed in future research. First, we used a sample of OA that was largely similar to the A group, except for age. While this was necessary to isolate the component of biological aging and to control for factors such as of computer literacy, required for this study, it reduces the generalizability of results among more vulnerable OA who may be more willing to benefit from the assistance of social robots [19]. In addition, and although the AO and A groups did significantly differ in age, the age gap between groups was not broad. Future research could consider a finer division of groups of age to explore if eye-gaze following declines also when it comes from a social robot. Second, the social nature of the current gaze cue remains unclear. For instance, Pepper’s head movement can be simply interpreted as a moving stimulus towards the correct answer. Future studies would benefit from including conditions where the signaling towards the correct ingredient is clearly social or non-social.

6 Conclusion

This study explored the influence of deictic gaze from a Pepper robot in two groups of age: adults (A) and older adults (OA). We found a facilitation effect of deictic gaze from a Pepper robot in all the participants independently of their age. These findings show that head movement representing deictic gaze is effective in terms of task performance. However, this facilitation effect was not significantly different between the age groups, which means that A do not benefit more than OA. Moreover, we found age-related differences in the effect of the robot’s deictic gaze on social perception. OA seem to be less reactive to deictic gaze than A when it comes to their report of anthropomorphism and discomfort caused by the robot.

Future research should add human and/or non-social controls to better inform the differences between the perception of human and robot gaze cues. In addition, the inclusion of non-valid cues would be useful to determine the role of trust and to explore the potential attentional costs. The results of this research line could ultimately be valuable in the design adaptive non-verbal cues from robots in HRI. This user-centered approach would allow a wider acceptance of social robots.

Acknowledgments. This work is funded by the EU Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 754285, by the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation, and by the RobWellproject (No RTI2018-095599-A-C22) funded by the Spanish *Ministerio de Ciencia, Innovación y Universidades*. We want to thank the universities that helped us with the sample recruiting process: Complutense University of Madrid, University Carlos III of Madrid, University of Murcia and University of Alicante. The authors thank Neziha Akalin and Estefanía Sánchez-Pastor for their support in the initial stages of this study.

References

1. Admoni, H., Scassellati, B.: Social eye gaze in human-robot interaction: A review. *Journal of Human-Robot Interaction* **6**(1), 25–63 (2017)
2. Admoni, H., Weng, T., Hayes, B., Scassellati, B.: Robot nonverbal behavior improves task performance in difficult collaborations. In: *Proc. ACM/IEEE Int. Conf. HRI (HRI'16)*. pp. 51–58. Christchurch, New Zealand (2016)
3. Andrist, S., Tan, X.Z., Gleicher, M., Mutlu, B.: Conversational gaze aversion for humanlike robots. In: *Proc. ACM/IEEE Int. Conf. HRI (HRI'14)*. pp. 25–32. Bielefeld, Germany (2014)
4. Baron-Cohen, S.: *Mindblindness: An Essay on Autism and Theory of Mind*. The MIT Press, Cambridge (1995)
5. Bartneck, C., Kulić, D., Croft, E., Zoghbi, S.: Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics* **1**(1), 71–81 (2009)
6. Carpinella, C.M., Wyman, A.B., Perez, M.A., Stroessner, S.J.: The robotic social attributes scale (RoSAS): Development and validation. In: *Proc. ACM/IEEE Int. Conf. HRI (HRI'17)*. pp. 254–262. New York, USA (2017)
7. Cañigual, R., Hamilton, A.F.C.: The role of eye gaze during natural social interactions in typical and autistic people. *Frontiers in Psychology* **10** (2019)
8. Faul, F., Erdfelder, E., Lang, A.G., Buchner, A.: G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods* **39**(2), 175–191 (2007)
9. Feingold-Polak, R., Elishay, A., Shahar, Y., Stein, M., Edan, Y., Levy-Tzedek, S.: Differences between young and old users when interacting with a humanoid robot: A qualitative usability study. *Paladyn* **9**(1), 183–192 (2018)
10. Finger, H., Goeke, C., Diekamp, D., Standvoß, K., König, P.: Labvanced: A unified javascript framework for online studies. In: *Proc. Int Conf. IC2S2 (IC2S2'17)*. Cologne, Germany (2017)
11. Harada, C.N., Love, M.C.N., Triebel, K.: Normal cognitive aging. *Clin Geriatr Med* **29**, 737 (2013)
12. Hart, S.G., Staveland, L.E.: Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Adv. Psychol* **52**, 139–183 (1988)
13. Kiilavuori, H., Sariola, V., Peltola, M.J., Hietanen, J.K.: Making eye contact with a robot: Psychophysiological responses to eye contact with a human and with a humanoid robot. *Biological Psychology* **158**, 107989 (1 2021)
14. Kontogiorgos, D., Pereira, A., Andersson, O., Koivisto, M., Rabal, E.G., Vartiainen, V., Gustafson, J.: The effects of anthropomorphism and non-verbal social behaviour in virtual assistants. In: *Proc. ACM Int. Conf. IVA (IVA'19)*. pp. 133–140. Paris, France (2019)
15. Morillo-Mendez, L., Mozos, O.M.: Towards human-based models of behaviour in social robots: Exploring age-related differences in the processing of gaze cues in human-robot interaction. In: *Starting Artificial Intelligence Researchers Symposium (STAIRS'20)*. Santiago de Compostela, Spain (2020)
16. Mutlu, B., Forlizzi, J., Hodgins, J.: A storytelling robot: Modeling and evaluation of human-like gaze behavior. In: *Proc. IEEE-RAS Int. Conf. HUMANOIDS*. pp. 518–523. Genoa, Italy (2006)
17. Mutlu, B., Shiwa, T., Kanda, T., Ishiguro, H., Hagita, N.: Footing in human-robot conversations. In: *Proc. ACM/IEEE Int. Conf. HRI (HRI'09)*. p. 61. La Jolla, California, USA (2009)

18. Mwangi, E., Barakova, E.I., Díaz-Boladeras, M., Mallofré, A.C., Rauterberg, M.: Directing attention through gaze hints improves task solving in human–humanoid interaction. *Int. J. Soc. Robot.* **10**(3), 343–355 (2018)
19. Pino, M., Boulay, M., Jouen, F., Rigaud, A.S.: ”Are we ready for robots that care for us?” Attitudes and opinions of older adults toward socially assistive robots. *Frontiers in Aging Neuroscience* **7** (2015)
20. Pu, L., Moyle, W., Jones, C., Todorovic, M.: The effectiveness of social robots for older adults: A systematic review and meta-analysis of randomized controlled studies. *Gerontologist* **59**, 37–51 (2019)
21. Slessor, G., Venturini, C., Bonny, E.J., Insch, P.M., Rokaszewicz, A., Finnerty, A.N.: Specificity of age-related differences in eye-gaze following: Evidence from social and nonsocial stimuli. *J. Gerontol. Series B* **71**, 11–22 (2016)
22. Zafrani, O., Nimrod, G.: Towards a holistic approach to studying human-robot interaction in later life. *Gerontologist* **59**, 26–36 (2019)